Cowles Foundation Paper 39c

PSYCHOMETRIKA—VOL. 15, NO. 2
JUNE, 1950

ON THE IDENTIFIABILITY OF PARAMETERS IN THURSTONE'S MULTIPLE FACTOR ANALYSIS*

OLAV REIERSØL
PURDUE UNIVERSITY†

In econometric literature a parameter in a theoretical model has been called identifiable if it can be uniquely determined in terms of the joint probability distribution of the observed variables. In this paper the identifiability of parameters in four different factor analysis models is considered. The last of these four models corresponds to Thurstone's factor analysis. In Sections 7 and 11, the possibility of a statistical testing of the models is discussed. Section 10 deals with the problem of actually determining the parameter r (the number of common factors) in terms of the probability distribution of the observed variables.

1. *Introduction*

Among the problems which Thurstone discusses in his books on multiple factor analysis are the questions of uniqueness of the communalities (11, 73-77, and 12, 307-311) and the uniqueness of the factor loadings (11, 155, and 12, 334). Instead of the term "uniqueness," which Thurstone uses, I shall in this paper use the term "identifiability," introduced by staff members of the Cowles Commission for Research in Economics. The question of the identifiability of parameters in theoretical models has been discussed by several writers in the econometric field before 1940, particularly by Ragnar Frisch. The main contributions to the formalization, generalization, and rigorous analysis of the problem were made by Haavelmo (3, Chapter 5), Koopmans and Rubin (6), and Hurwicz (4), all writing in the econometric field.

The theoretical model considered in this paper has some analogy to a particular case of the econometric model considered by Koopmans and Rubin in the paper just referred to. The fundamental equation (2.1) in the present paper has its counterpart in the reduced form in the econometric model, but in the econometric model the variables X are observed variables, while in the psychometric model they are latent variables. Because of this difference the analogy men-

*This article will be included in Cowles Commission Paper, New Series, No. 39.

†Present address: University of Oslo, Oslo, Norway.

tioned is of no help in the discussion of identifiability in the factor analysis model. Theorem 5.1, however, brings in another analogy between the two models. Again there is the difference that the identifiability of the matrix A in factor analysis depends upon assumptions about a certain minimum number of zero elements in each column of A, the locations of which are not known *a priori*, while in the econometric model the identifiability of the corresponding matrix depends on assumptions about given elements of the matrix being zero, or about linear relationships between given elements of the matrix. Because of this difference the discussion of the identifiability is more complicated in the factor analysis model than in the econometric model.*

2. *The theoretical model of multiple factor analysis*

We shall use the following notation:

$s_j =$ score in test j,
$x_m =$ score in common factor m,
$a_{jm} =$ factor loading of common factor m in test j,
$u_j =$ unique part of s_j (i.e. specific factor + error),
$n =$ number of tests, and
$r =$ number of common factors.

We shall introduce the following row vectors:

$$S = [s_1, s_2, \cdots, s_n],$$
$$X = [x_1, x_2, \cdots, x_r], \text{ and}$$
$$U = [u_1, u_2, \cdots, u_n];$$

and the matrix:

$$A = \begin{bmatrix} a_{11}, a_{12}, \cdots, a_{1r} \\ a_{21}, a_{22}, \cdots, a_{2r} \\ \cdot \; \cdot \; \cdot \; \cdot \; \cdot \; \cdot \; \cdot \; \cdot \; \cdot \; \cdot \\ a_{n1}, a_{n2}, \cdots, a_{nr} \end{bmatrix}$$

M, Q, and D will be used to denote the population covariance matrices of the vector variables S, X, and U, respectively. The transpose of a matrix A will be denoted by A'. The diagonal elements of the matrix $M-D$ will be called *communalities,* and the diagonal elements of the matrix D will be called *uniquenesses.* Thurstone uses these terms after standardization of the variables S.

*A more detailed comparison between the two models is given in a joint paper by Koopmans and myself (7). I am also indebted to Professor Koopmans for valuable suggestions and criticisms regarding the present paper.

The variables S are the only observed variables. The variables X and U are not observed and will be called *latent variables.*

Our model will be given by the following specifications, all of which refer to the hypothetical population, not to a sample. For the population mean (mathematical expectation) we shall use the usual symbol E .

$$S' = AX' + U'. \tag{2.1}$$

$$E(X) = 0 \quad \text{and} \quad E(U) = 0\,. \tag{2.2}$$

The set of variables X is stochastically independent of the set of variable U . (2.3)

D is diagonal and different from 0 .* (2.4)

X and U are normally distributed. (2.5)

A has rank r . (2.6)

Q is nonsingular. (2.7)

Each s_i is correlated with at least one of the other s's. (2.8)

r is the smallest number of variables X which is compatible with the joint probability distribution of the observed variables and specifications (2.1) through (2.8) (2.9)

Each column of A contains at least r zeroes. (2.10)

A normalization rule fixing the units of the variables X and a rule fixing the order of the columns of A . (2.11)

Consider the matrix consisting of all the rows of A which have a zero in the mth column. Let this matrix be denoted by A_m . Let the number of rows in the matrix A_m be p_m . Let A_{mi} denote the submatrix of A_m constructed by deleting the ith row of A_m . Using these notations we shall formulate the final specification of our model.

The rank of each of the matrices A_{mi} ($m = 1, 2, \cdots, r; i = 1, 2, \cdots, p_m$) is $r - 1$. (2.12)

Specification (2.2) does not represent any loss of generality.

*A matrix (or vector) is said to be equal to zero when each of its elements is equal to zero. It is said to be different from zero if at least one of its elements is different from zero.

Specifications (2.4) and (2.5) imply that the variables U are mutually independent. The assumption that D is different from zero is made for convenience in order to avoid discussion of a case of no practical interest. It is equivalent to the assumption that M is non-singular. Specification (2.8) implies that no row of the matrix A contains only zero elements.

The model considered represents an interpretation of Thurstone's multiple factor analysis. This presentation of the specifications is, however, somewhat different from that of Thurstone, partly because we have chosen the specifications in such a way that a rigorous discussion of identifiability should not become too complicated, and partly because we have tried to formulate specifications which are in a sense minimum conditions for identifiability.

We have included in our model Thurstone's assumption of simple structure, but not the assumption that the factor loadings are non-negative. Thurstone says about the postulate of non-negative factor loadings (12, 341), "But it must be understood that this is a postulate of much psychological investigation in a particular field and that it does not constitute a restriction upon the principle of simple structure, which is applicable either with or without the special restriction that the factor loadings shall be positive or zero."

If all parameters A, Q, and D in our model are given numerically, we shall talk about a *structure*. The structure includes completely specified distributions of the latent variables and a set of equations with numerically given coefficients connecting the observed and latent variables. A structure thus is a particular realization of the model, and the model is the set of all structures compatible with the given specifications.

From Equations (2.1) and (2.2) we obtain

$$E(S) = 0 \tag{2.13}$$

and

$$M = E(S'S) = E[(AX' + U')(XA' + U)] = AE(X'X)A' + AE(X'U) + E(U'X)A' + E(U'U).$$

From (2.3) this reduces to

$$M = AQA' + D. \tag{2.14}$$

From (2.1) and (2.5) we conclude that the variables S are jointly normally distributed.

A given structure $T = \{A, Q, D\}$ generates one and only one joint probability distribution $P(S)$ of the observed variables S. On the other hand there may possibly be several structures T generating the same distribution $P(S)$. If two or more structures generate the same joint probability distribution of the observed variables, the structures are said to be *equivalent.* If a parameter has the same value in all equivalent structures it is said to be *identifiable.* In other words, a parameter is identifiable if it can be uniquely determined from a knowledge of the joint probability distribution of the observed variables. If a parameter is not identifiable, no consistent estimate of the parameter will exist.

3. Consideration of the model defined by specifications (2.1) through (2.8)

For this model we shall in the following use the notation Model (2.1-8). We shall study this model as a preliminary step in the study of the complete model given in Section 2.

Let a distribution $P(S)$ with a covariance matrix M be given, and suppose that we have a set of three matrices A, Q, and D satisfying Equation (2.14) and the following conditions:

A has the same rank as its number of columns, (3.1)

Q is a symmetric and positive definite, (3.2)

D is diagonal, non-negative, and different from 0. (3.3)

Let X and U be two sets of normally distributed variables with the joint probability density

$$\frac{\sqrt{(\det Q)\,(\det D)}}{(2\pi)^{\frac{n-r}{2}}}\, e^{-\frac{1}{2}(XQX'+UDU')},$$

where r is the order of the matrix Q. These variables satisfy Specifications (2.2) through (2.7), and when we define the set of variables $\bar{S}$ by $\bar{S}' = AX' + U'$, the variables $\bar{S}$ will have the same joint distribution as the given variables S. Hence we have the following theorem:

Theorem 3.1 The set of structures equivalent in Model (2.1-8) and generating a distribution P(S) *with covariance matrix* M *is given by the set of all matrices* A, Q, *and* D *satisfying Equation (2.14) and Conditions (3.1) through (3.3).*

When the matrix Q is positive definite, the matrix AQA' is positive semidefinite (when $r < n$) or positive definite (when $r = n$). Hence we have:

$$M - D \text{ is positive definite or semidefinite.} \tag{3.4}$$

Suppose inversely that (3.4) holds good. Now any positive definite or positive semidefinite matrix can be transformed by a cogredient transformation to any other positive definite or semidefinite matrix of the same rank. If the rank of $M—D$ is r, we can therefore find a non-singular matrix H_1 such that

$$M - D = H_1 \begin{bmatrix} Q & 0 \\ 0 & 0 \end{bmatrix} H_1',$$

where Q is a symmetric, r-rowed, positive definite matrix. Let H_2 denote the matrix consisting of the r first columns of H_1 and let H_3 denote the matrix consisting of the last $n - r$ columns of H_1. Then we can write

$$M - D = [H_2 H_3] \begin{bmatrix} Q & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} H_2' \\ H_3' \end{bmatrix},$$

which reduces to

$$M - D = H_2 Q H_2'.$$

This gives a solution of (2.14) with $A = H_2$, and it is immediately seen that this solution satisfies Conditions (3.1) through (3.3). Hence we have

Theorem 3.2. Let the covariance matrix M *and the order* r *of the matrix* Q *be given. Then a necessary and sufficient condition for the existence of a solution of Equation (2.14) satisfying Conditions (3.1) through (3.3) is that there exists a matrix* D *such that Conditions (3.3) and (3.4) are satisfied and the rank* M — D *is* r. *When a solution exists, there exist infinitely many solutions, since the matrix* Q *can be any symmetric,* r*-rowed, and positive definite matrix.*

Next we shall prove

Theorem 3.3. Suppose that Model (2.1-8) contains a structure T *where* r = s < n *and where* D *is non-singular. Then the model contains an infinity of structures equivalent to* T *where* r = s + 1 *and where* D *is non-singular.*

Let A, Q, D be a structure where $r = s$ and the determinant of $D \neq 0$. Then $M - D$ is of rank s, and it contains at least one

s-rowed principal minor which is different from zero. Let P be such a minor and let $m_{ii} - d_i$ be a diagonal element not contained in P. Let us consider the $(s + 1)$-rowed minor P_1 which we obtain when adding the ith row and the ith column of $M - D$ to the minor P. Let us change d_i to d_i^*, where $0 < d_i^* < d_i$. This change will increase the value of P_1 and, since it was zero before the change, it must be positive after the change. Let D^* be a diagonal matrix which we get from D by changing the single element d_i to d_i^*. Then D^* is non-negative and non-singular and $M - D^*$ is positive semidefinite and of rank $s + 1$. Using Theorems 3.1 and 3.2, Theorem 3.3 follows.

Let the diagonal elements $d_1, d_2, \cdots, d_n$ of the matrix D be interpreted as Cartesian coordinates in an n-dimensional space which we shall call D-space, and let the covariance matrix M be fixed and non-singular. Then there exists an n-dimensional region in D-space where D is non-negative and different from zero and where $M - D$ is positive definite, and there exists an $(n - 1)$-dimensional region in D-space where D is non-negative and different from zero and where $M - D$ is positive semidefinite and of rank $n - 1$. (Cf. 9, p. 15). Hence there exist structures in Model (2.1-8) where $r = n$, and structures where $r = n - 1$, all of them generating the given covariance matrix M. Hence r is not identifiable in this model, but may have at least two values, n and $n - 1$. And for each of these values of r there exists an infinity of possible matrices D. Again for each possible D there exists an infinity of possible matrices A and Q. We are thus very far from identifiability in this model.

If we drop the specification that the matrix A has rank r, any value of r which is greater than $n - 2$ is possible. The lack of identifiability in a model of this kind was pointed out in 1919 by G. H. Thomson (10), who gave numerical examples of different structures generating the same covariance matrix of the observed variables.

4. *Identifiability of the uniquenesses in a model defined by Specifications (2.1) through (2.9)*

We may reformulate Specification (2.9) as follows: Consider the set of all equivalent structures in Model (2.1-8) which generate a given distribution $P(S)$. r is the least number of variables X occurring in any of these structures. A third way of formulating Specification (2.9) is the following: For a given covariance matrix M

consider the set of all solutions of Equation (2.14) satisfying conditions (3.1)-(3.3). r is the least value of the order of Q occurring in any solution in this set. A fourth way of formulating the specification will be: For a given covariance matrix M consider the set of all matrices D such that (3.3) and (3.4) hold. r is the minimum rank of $M - D$ within this set.

It follows immediately from any of these formulations that r is always identifiable in Model (2.1-9), and from the considerations in Section 3 it follows that $r \leqq n - 1$.

We shall next consider the identifiability of the uniquenesses, which are the diagonal elements d_i of the matrix D. Walter Ledermann has shown (8) that for the determination of the d's we have

$$k_r = \tfrac{1}{2}\,(n - r)\,(n - r + 1)$$

equations, which are generally independent. We shall make this statement more precise. Let the elements of the matrix M be interpreted as Cartesian coordinates in a $\dfrac{k(k+1)}{2}$-dimensional space, which we shall call M-space. The k_r equations will be independent at all points of M-space except for a subspace of lower dimensionality.

We have $k_r \begin{smallmatrix} < \\ = \\ > \end{smallmatrix} n$ according as $r \begin{smallmatrix} > \\ = \\ < \end{smallmatrix} r_n$, where

$$r_n = \tfrac{1}{2}\,(2n + 1 - \sqrt{8n + 1}).$$

The k_r equations are not linear. In the case when $r = r_n$ we shall therefore usually not have one solution only, (except in the case when $n = 3$, $r = 1$) but a finite number of solutions. In order to be admissible these solutions must satisfy Conditions (3.3) and (3.4). These conditions may or may not exclude all solutions except one (15).

When $r > r_n$ we shall generally expect D to be non-identifiable, but it may be identifiable for some sets of values of the elements of M. When $r < r_n$ we shall expect D to be identifiable, but it will fail to be identifiable for some sets of values of the elements of M.

Even though we do not know much about the exact conditions for identifiability of the uniquenesses, we may state that the identifiability is in principle subject to statistical test. It has been shown that the identifiability of parameters in any model is in principle subject to statistical test (7).

5. *The set of equivalent structures in Model (2.1-9)*

We shall find the set of equivalent structures in the case when D is identifiable. Two equivalent structures must then have the same D. Let the structures $T = \{A, Q, D\}$ and $T^* = \{A^*, Q^*, D\}$ be equivalent. Then

$$M - D = AQA' = A^* Q^* A^{*\prime}. \tag{5.1}$$

We can write this equation in the form

$$\begin{bmatrix} A_I \\ A_{II} \end{bmatrix} Q[A_I' A_{II}'] = \begin{bmatrix} A_I^* \\ A_{II}^* \end{bmatrix} Q^*[A_I^{*\prime} A_{II}^{*\prime}], \tag{5.2}$$

where A_I and A_I^* represent the upper square parts of A and A^*, respectively. Since A is of rank r we can without loss of generality assume that A_I is non-singular. From (5.2) we get 4 equations between the submatrices. We shall use two of these equations,

$$A_I QA_I' = A_I^* Q^* A^{*\prime} \tag{5.3}$$

and

$$A_{II} QA_I' = A_{II}^* Q^* A_I^{*\prime}. \tag{5.4}$$

Setting

$$V = A_I^{-1} A_I^*, \tag{5.5}$$

V is a square, r-rowed, non-singular matrix, and we get from (5.3)

$$Q = VQ^* V'. \tag{5.6}$$

From (5.5) we get

$$A_I^* = A_I V. \tag{5.7}$$

Using (5.6) and (5.7) we obtain from (5.4)

$$A_{II}^* = A_{II} V. \tag{5.8}$$

Equations (5.7) and (5.8) can be written together as one single matrix equation

$$A^* = AV. \tag{5.9}$$

We have now shown that (5.6) and (5.9) are necessary conditions for the equivalence of the structures T and T^*. Inversely, suppose that the structure $T = \{A, Q, D\}$ generates the covariance matrix M, and let A^* and Q^* be defined by (5.6) and (5.9), where V is any square r-rowed, non-singular matrix. Then it follows that (5.1) holds good and that A^* and Q^* satisfy conditions (3.1) and (3.2). Hence we have the following theorem:

Theorem 5.1: The set of all structures $\{A^*, Q^*, D\}$ *equivalent in Model (2.1-9) to the structure* $\{A, Q, D\}$ *is given by the set of all matrices* $A^* = AV$ *and* $Q^* = V^{-1}Q(V')^{-1}$, *where* V *is any square, r-rowed and non-singular matrix.*

Among the matrices A^* there may possibly be some matrices where one or more columns contain only one non-zero element. This would mean that the factor occurs only in a single test and will therefore not be a common factor. We have, however, not excluded such a case by our specifications. Strictly speaking the x's need not represent common factors. It should be noted, however, that no x can be independent of all the other x's and at the same time occur in one test only. For, suppose that the factor x_m occurs in the jth test only. Then we may include $a_{jm}x_m$ in u_j instead of in $s_j - u_j$. After this change our specifications will still hold good with r replaced by $r - 1$. But this contradicts the specification defining r and is therefore impossible.

6. *Identifiability in Model (2.1-11)*

Specifications (2.1) through (2.10) are never sufficient for the identifiability of A, for if we have one matrix A satisfying these specifications, they will still be satisfied if we multiply each column by an arbitrary factor and perform an arbitrary permutation of the columns. In order to take care of this indeterminancy we have introduced Specification (2.11). The normalization rule fixing the units of the variables X may be formulated in various way in terms of the elements of the matrix Q or in terms of the elements of the matrix A. We shall give three examples of possible formulations.

The variance q_{mm} of x_m is equal to 1.00 $(m = 1, 2, \cdots, r)$. (6.1)

$$\sum_{j=1}^{n} a_{mj} = 1, \quad (m = 1, 2, \cdots, r). \tag{6.2}$$

$$\sum_{j=1}^{n} |a_{mj}| = 1, \quad (m = 1, 2, \cdots, r). \tag{6.3}$$

Also the rule fixing the order of the columns of A can be formulated in various ways.

The following theorem is evident:

Theorem 6.1. A necessary and sufficient condition for A *to be identifiable in Model (2.1-11) is that all matrices* A^* *equivalent to* A *in Model (2.1-10) can be written in the form* $A\bar{V}$, *where* $\bar{V}$ *is a square non-singular matrix which contains exactly one non-zero element in each row and exactly one non-zero element in each column.*

The addition of Specification (2.10) may possibly make the matrix D identifiable even if it is not identifiable in Model (2.1-9). We shall, however, not enter into any discussion of this question, but shall confine the discussion to the identifiability of A in the case when D is identifiable in Model (2.1-9).

We shall now introduce a geometrical representation of the matrix A. (Cf. 12, Chapter 3). Before we do that we shall define a few concepts of multidimensional geometry. The set of points in n-dimensional space whose Cartesian coordinates satisfy a single linear equation will be called a *hyperplane.* The intersection of two or more hyperplanes will be called a *flat.* If it is of p dimensions, it will be called a *p-flat.* A one-flat is a straight line, a two-flat is a plane, and an $(n-1)$-flat is a hyperplane. A figure bounded by hyperplanes will be called a *polytope.* The simplest polytope in n-dimensional space is bounded by $n+1$ hyperplanes. It is called a *simplex.* A simplex in two-dimensional space is a triangle, in three-dimensional space a tetrahedron.

We shall now consider the geometrical representation already announced. Let the ith row of the matrix A be denoted by a_i. Let us choose a Cartesian coordinate system C in r-dimensional space, and let g_i be a vector whose components in the coordinate system C are the elements of the row a_i. Let $e_1, e_2, \cdots, e_r$ be unit vectors along the coordinate axes in the system C. Then

$$g_i = \sum_{j=1}^{r} a_{ij}\, e_j. \tag{6.4}$$

Let us next choose a new Cartesian coordinate system C^*, and let $e_1^*, e_2^*, \cdots, e_r^*$ denote unit vectors along the axes in the system C. Let v_{jk} be the kth component of e_j in the coordinate system C^*, so that

$$e_j = \sum_{k=1}^{r} v_{jk}\, e_k^*. \tag{6.5}$$

Let a_{ik}^* be the kth component of g_i in the system C, so that

$$g_i = \sum_{k=1}^{r} a_{ik}^*\, e_k^*. \tag{6.6}$$

Substituting (6.5) in (6.4) we obtain

$$a_{ik}^* = \sum_{j=1}^{r} a_{ij}\, v_{jk}, \tag{6.7}$$

which can be written in matrix form

$$A^* = AV. \tag{6.8}$$

To each transformation V of the coordinate system corresponds one and only one matrix A^* in the set of matrices which are equivalent in Model (2.1-9) to the matrix A. Inversely, if a matrix A^* equivalent to A is given, V is uniquely determined by (5.5), and we can find a coordinate system C^* such that the ith row of a_i of A gives the components of the vector g_i in the coordinate system C^*, for $i = 1, 2, \cdots, n$. The axes of the coordinate system C^* are uniquely determined by the equations

$$e_j^* = \sum_{k=1}^{r} v^{jk}\, e_k,$$

where v^{jk} are the elements of the matrix V^{-1}. This gives a possible coordinate system, since we do not assume that the coordinate systems are orthogonal. There is thus a one-to-one correspondence between the set of all possible Matrices A^* equivalent in Model (2.1-9) to a given matrix A, and the set of all possible Cartesian coordinate systems C with the same origin as C. Specification (2.10) can now be reformulated thus:

Each coordinate hyperplane contains at least r row vectors of the matrix A. (2.10b)

The problem of the identifiability of A, given that of D, is now seen to be equivalent to the question of whether the coordinate hyperplanes are uniquely determined by Condition (2.10b). The length of the units along the coordinate axes and the numbering of the coordinate axes are not determined by (2.10b), but will be determined by (2.11).

It it possible to get a geometric representation where the number of dimensions is reduced by one. This can be done by intersecting all test vectors g_i or their extensions forward or backward by a hyperplane H, which does not pass through the origin. If we include the points at infinity in this hyperplane there is a one-to-one correspondence between the test vectors g_i and the intersection points h_i in the hyperplane. These intersection points will be called test points. The intersections between the hyperplane H and the coordinate hyperplanes in r-dimensional space will again be called coordinate hyperplanes in $(r-1)$-dimensional space. They will be bounding hyperplanes of an $(r-1)$-dimensional simplex, which we shall call the fundamental simplex. The intersection between two or more coordinate hyperplanes will be called a coordinate p-flat if it is of dimensionality p. We can now give a second reformulation of specification (2.10).

$$\text{Each coordinate hyperplane in the space } H \text{ contains at least } r \text{ of the test points } h_i . \tag{2.10c}$$

The problem of the identifiability of A is seen to be equivalent to the question of whether the coordinate hyperplanes in the space H are uniquely determined by Condition (2.10c).

We shall first show that Specification (2.10) in a certain sense is a minimum requirement for identifiability of the matrix A. This will be expressed in the following theorem:

Theorem 6.2 Let the number r *in Specification* (2.10) *be replaced by* r — 1. *Then the matrix* A *is not identifiable.*

Proof: Suppose as before that the upper square part A_I of A is non-singular and consider the matrix

$$C = \begin{bmatrix} A_I \\ A_{II} \end{bmatrix} A_I^{-1} = \begin{bmatrix} I \\ C_{II} \end{bmatrix},$$

where $C_{II} = A_{II}A_I^{-1}$. From Specification (2.8) follows that each row of A must be different from zero. Since A_I^{-1} is non-singular, also each row of C must be different from zero. Since $r \leqq n-1$, the matrix C_{II} must contain at least one row. Let $c_1, c_2, \cdots, c_r$ be the elements of the first row of C_{II}. Without loss of generality we can assume that $c_1 \neq 0$. Post-multiplying the matrix C by the matrix

$$V_1 = \begin{bmatrix} \frac{1}{c_1} & -\frac{c_2}{c_1} & \cdots & -\frac{c_r}{c_1} \\ 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

we get the matrix

$$C^* = \begin{bmatrix} \frac{1}{c_1} & -\frac{c_2}{c_1} & \cdot\cdot\cdot\cdot & -\frac{c_r}{c_1} \\ 0 & 1 & \cdot\cdot\cdot\cdot\cdot & 0 \\ \cdot & \cdot & \cdot\cdot\cdot\cdot\cdot & \cdot \\ 0 & 0 & \cdot\cdot\cdot\cdot\cdot & 1 \\ 1 & 0 & \cdot\cdot\cdot\cdot\cdot & 0 \end{bmatrix}$$

This matrix has at least $r-1$ zeros in each column. If at least one of the elements $c_2, c_3, \cdots, c_r$ is different from zero, the transformation V_1 is not of the form $\bar{V}$. The only case which is still in doubt is when c_1 is the only element in the first row of C_{II} which is different from zero.

In this case we postmultiply C by the following matrix:

$$V_2 = \begin{bmatrix} 1 & 0 & \cdot\cdot\cdot\cdot\cdot & 0 & 0 \\ 0 & 1 & \cdot\cdot\cdot\cdot\cdot & 0 & 0 \\ \cdot & \cdot & \cdot\cdot\cdot\cdot\cdot & \cdot & \cdot \\ 0 & 0 & \cdot\cdot\cdot\cdot\cdot & 1 & 0 \\ 0 & p_2 & \cdot\cdot\cdot\cdot\cdot & p_{r-1} & p_r \end{bmatrix}$$

where $p_2, \cdots, p_{r-1}, p_r$ are arbitrary. We then get the matrix

$$C^{**} = \begin{bmatrix} 1 & 0 & \cdot\cdot\cdot\cdot\cdot & 0 & 0 \\ 0 & 1 & \cdot\cdot\cdot\cdot\cdot & 0 & 0 \\ \cdot & \cdot & \cdot\cdot\cdot\cdot\cdot & \cdot & \cdot \\ 0 & 0 & \cdot\cdot\cdot\cdot\cdot & 1 & 0 \\ 0 & p_2 & \cdot\cdot\cdot\cdot\cdot & p_{r-1} & p_r \\ C_1 & 0 & \cdot\cdot\cdot\cdot\cdot & 0 & 0 \\ \cdot & \cdot & \cdot\cdot\cdot\cdot\cdot & \cdot & \cdot \end{bmatrix}$$

and in this matrix again each column contains at least $r-1$ zeros. Choosing both p_r and at least one of the other p's different from zero, the transformation matrix V_2 is non-singular and not of the type $\bar{V}$. This completes the proof of Theorem 6.2.

Theorem 6.3. When D *is identifiable in Model (2.1-9), a necessary and sufficient condition for* A *to be identifiable in Model (2.1-11) is that any square* r*-rowed minor of* A *which is of rank* r − 1 *is contained in one of the matrices* A_m.

This theorem may be reformulated in terms of the geometric representation in the space H as follows:

Theorem 6.3b. When D *is identifiable in Model (2.1-9), a necessary and sufficient condition for* A *to be identifiable in Model (2.1-11) is that any hyperplane containing at least* r *test points is a coordinate hyperplane.*

When the condition is satisfied the r coordinate hyperplanes are uniquely determined as the r hyperplanes each of which contains at least r test points. On the other hand, if there exists a hyperplane which is not a coordinate hyperplane, but which contains at least r test points, then there exists an equivalent structure where this hyperplane is a coordinate hyperplane. This proves the theorem.

Suppose that a coordinate $(r-3)$-flat contains $r-1$ test points. Then any hyperplane H' through this $(r-3)$-flat and any other test point will contain at least r test points. Suppose that there is at least one test point not lying in one of the two coordinate hyperplanes whose intersection is the $(r-3)$-flat. Then a hyperplane H' through this test point and the $(r-3)$-flat will contain r test points, and H' will be different from any of the coordinate hyperplanes. Suppose next that a coordinate $(r-3)$-flat contains r or more test points. Then any hyperplanes through the $(r-3)$-flat will contain r test points. In this case we thus have an infinity of hyperplanes H' each of which contains at least r test points and such that each H' is different from any coordinate hyperplane.

The condition of Theorem 6.3 will therefore generally exclude the case of a coordinate $(r-3)$-flat containing $r-1$ test points and it will always exclude the case of a coordinate $(r-3)$-flat containing r or more test points. But according to Thurstone these excluded cases are what we may expect in practice, and they are even desirable (12, 335, Condition 4). The results of the present section are therefore not practically useful. We want a model where the coordinate hyperplanes are characterized not only by the requirement that each of them shall contain at least r points, since we shall expect many other hyperplanes to contain r or more points. A natural assumption is that each coordinate hyperplane shall be *overdetermined* by the points in it, i.e., if we take away any single point in the coordinate hyperplane, the remaining points will still determine the hyperplane. This is just what we have expressed algebraically in Specification (2.12). But before we consider the problem of identifiability of the parameters in the complete model we shall consider the possibility of testing statistically the specifications.

7. On the possibility of testing the specifications

If any joint probability distribution of the observed variables can be generated from a structure belonging to the model we are considering, the model cannot be tested statistically. Let us first consider Model (2.1-9). We can test Specification (2.5), taking for granted that Specifications (2.1) and (2.3) hold good. Specification (2.8) can also be tested, since it is an assumption about the observed variables. No further testing of the model is possible, since any covariance matrix M satisfying Specification (2.8) can be generated by a structure in the model.

In Model (2.1-11) it is in principle possible to test Specification (2.10) assuming that Specifications (2.1) through (2.9) hold good. And in Model (2.1-12) it is in principle possible to test Specifications (2.10) and (2.12), taking for granted that Specifications (2.1) through (2.9) hold good. We shall consider the latter test only. The set of admissible hypotheses is now given by all structures satisfying Specifications (2.1) through (2.9), and the hypothesis to be tested includes all structures satisfying Specifications (2.1) through (2.12).

Let the matrix A belong to a structure in Model (2.1-9) and let B be any matrix of the form $B = AV$, where V is square and non-singular. Let us consider the matrices $A_1, A_2, \cdots, A_r$ considered before, and let B_m be a submatrix of B consisting of the rows of B which have the same position in B as the rows of A_m have in A. Then $B_m = A_m V$, and since V is non-singular, the rank of B_m must be the same as the rank of A_m. Let B_{mi} denote the submatrix of B which we get when we delete the ith row of B_m. Then $B_{mi} = A_{mi} V$, hence the rank of B_{mi} is the same as the rank of A_{mi}.

Suppose now that Specifications (2.10) and (2.12) are satisfied for the matrix A. Since A_m contains a column consisting of zeros only, its rank cannot be greater than $r - 1$, and since it contains submatrices of rank $r - 1$, its rank must be exactly $r - 1$. Hence we get

$$\text{The rank of each } B_m \text{ is } r - 1 . \tag{7.1}$$

$$\text{The rank of each } B_{mi} \text{ is } r - 1 . \tag{7.2}$$

The matrix A_m is supposed to contain all rows with zeros in the mth column. Hence, if we add to A_m a row of A which is not contained in A_m, the resulting matrix must be of rank r, and from this we again conclude:

The addition to B_m of a row of B not contained in B_m increases the rank to r. (7.3)

In addition to the matrices $B_1, B_2, \cdots, B_r$, there may possibly exist other submatrices of B with r columns, satisfying Conditions (7.1) through (7.3). If such matrices exist we shall denote them by $B_{r+1}, \cdots, B_k$.

Setting $W = V^{-1}$, W is a square, r-rowed, non-singular matrix and $A = BW$. Let W_m denote the mth column of W. Since the mth column of A_m consists of zeros only, we have $B_m W_m = 0$. Since the columns of W are linearly independent we can state

The set of right nullspace vectors* of the matrices $B_1, B_2, \cdots, B_k$ contains r linearly independent vectors. (7.4)

Suppose inversely that a matrix B satisfies Conditions (7.1) through (7.4). Among the right nullspace vectors W_i of the matrices B_i we choose r which are linearly independent, and we form a matrix W the columns of which are these r nullspace vectors. Then it is evident that the matrix $A = BW$ satisfies Specifications (2.10) and (2.12). We shall formulate our results in the following theorem:

Theorem 7.1. Let B *be any matrix* AV, *where* V *is a square, non-singular matrix. A necessary and sufficient condition for the matrix* A *to satisfy Specifications (2.10) and (2.12) is that the matrix* B *satisfies Conditions (7.1)-(7.4).*

Suppose now that the population covariance matrix M is known, and that the matrix D is identifiable. Then the matrix $M - D$ can be determined and we can find a matrix B such that $BB' = M - D$, for instance by the diagonal method of factoring (12, 101). Next we can examine if the matrix B satisfies Conditions (7.1) through (7.4). If it does satisfy these conditions, we know that there exists a structure in Model (2.1-12) generating a distribution $P(S)$ with a covariance matrix M. If B does not satisfy Conditions (7.1) through (7.4), Model (2.1-12) must be rejected. If Specifications (2.1) through (2.9) are taken for granted we shall have to reject Specifications (2.10) and (2.12) or at least Specification (2.12).

So far we have discussed what conclusions we can draw about the specifications when we know the population covariance matrix.

*A right nullspace vector of a matrix is a vector which vanishes after a premultiplication by the matrix.

When we do not know the population covariance matrix, as is the case in practice, but know only a sample, it follows, however, from the foregoing considerations that it is possible, at least in point of principle, to test Specifications (2.10) and (2.12).

8. Reformulation of Specifications (2.12) in terms of assumptions about non-zero elements

Instead of Specification (2.12), Thurstone makes assumptions about non-zero elements in the matrix A (**11**, 156 and **12**, 335). We shall reformulate Specification (2.12) in a similar form.

Let matrix B be considered as a function of some of its elements while the other elements are constants. Let the elements of B which we regard as variables be denoted by $z_1, z_2, \cdots, z_p$ and let Z denote the vector

$$Z = [z_1, z_2, \cdots, z_p] .$$

The matrix B regarded as a function of the variables Z will be denoted by $B(Z)$. Any minor determinant in B will be a polynomial in the variables Z or identically equal to zero. If it is not identically equal to zero it will be equal to zero only in a set of measure zero in Z-space. The rank of the matrix B will be a function of Z. The maximum of the rank B when Z varies will be called the rank of the matrix function $B(Z)$. Evidently this rank is the rank of B in the whole Z-space, except for a set of measure zero.

Let Y be a vector whose components are all elements of B which are not components of Z. So far we have considered Y as constant. Let us now consider the minimum rank of $B(Z)$ when Y varies over all possible values. This minimum will be called the minimum rank of $B(Z)$.

Lemma 8.1. B(Z) *attains its minimum rank when* $\mathrm{Y}=0$.

Proof: Any minor determinant of B is a polynomial in the variables Z. The coefficient of each term in the polynomial is either *1* or *—1* or a homogeneous function of the variables Y of degree one or higher degree. Hence, if a coefficient in the polynomial vanishes for some value of Y, it vanishes when $Y = 0$. From this follows that if a minor determinant of B vanishes identically in Z for some value of Y, it will vanish identically in Z when $Y = 0$. This proves the lemma.

Lemma 8.2. Let B *be a matrix with* r *columns and a number of rows which is at least equal to* r . *When* $\mathrm{Y} = 0$, *a necessary and sufficient condition for the rank of* B(Z) *to be equal to* r *is that the* *following condition is satisfied for* q $= 1, 2, \cdots$, r: *For each set of* q *columns of* B *there exist at least* q *rows, each of which has a variable* z *in at least one of the* q *colunms.*

Proof: We shall first prove the necessity of the condition. Suppose that the condition is not satisfied. Then there exists for some q a set of q columns of B which contain elements of Z in at most q—1 rows. Any r-rowed minor of B may be expanded in a Laplace expansion according to the q columns. Since all q-rowed determinants in these q columns are zero, each r-rowed minor of B must be zero identically in the variables Z. Hence the rank of $B(Z)$ is less than r. This proves the necessity of the condition in Lemma 8.2.

We shall next prove the sufficiency of the condition. We shall prove this by induction. The condition is evidently sufficient when $r = 1$. The sufficiency statement will therefore be proved in the general case if we can prove that it is true for a matrix B with r columns provided that it is true for any matrix with less than r columns. We thus have to prove that if the condition is satisfied for a matrix B with r columns and if Lemma 8.2 is true for any matrix with less than r columns, then the rank of $B(Z)$ must be r. In order to prove this we shall assume that the rank of B is $r - 1$ and show that this leads to a contradiction.

Let us consider the submatrix of B obtained by deleting the rth column of B. According to our assumptions this submatrix is of rank $r - 1$, and there exists a product $b_{i_1,1}\, b_{i_2,2} \cdots b_{i_{r-1},r-1}$ where each factor is a z. By an interchange of rows if necessary we can obtain that $i_1 = 1$, $i_2 = 2$, $\cdots i_{r-1} = r - 1$. Let us next consider the product $b_{11}\, b_{22} \cdots b_{r-1,r-1}\, b_{ir}$ where $i \geqq r$. Since the rank of B is $r - 1$ by assumption, we must have $b_{ir} = 0$ when $i \geqq r$. We now have the following picture of the matrix B.

$$\begin{matrix} z \\ & z \\ & & z \\ & & & z \\ & & & & \cdot \\ & & & & & \cdot \\ & & & & & & \cdot \\ & & & & & & & z \\ & & & & & & & & z \\ & & & & & & & & & 0 \\ & & & & & & & & & 0 \\ & & & & & & & & & \cdot \\ & & & & & & & & & \cdot \\ & & & & & & & & & \cdot \\ & & & & & & & & & \cdot \\ & & & & & & & & & 0 \end{matrix}$$

Column r contains at least one z, and since $b_{ir} = 0$ when $i \geqq r$, a z must occur in one of first $r - 1$ rows. If $b_{r-1,r}$ is not a z we can always by an interchange of rows and the corresponding interchange of columns obtain that $b_{r-1,r}$ will be a z, while the elements in the main diagonal which are known to be z's still will be situated in the main diagonal. Consider now the product

$$b_{11}\, b_{22} \cdots b_{r-2,r-2}\, b_{r-1,r}\, b_{i,r-1}$$

where $i \geqq r$. Since B is of rank $r - 1$, all these products must be zero; hence we conclude that $b_{i,r-1} = 0$ when $i \geqq r$, since all other factors in the product above are known to be z's. Columns $r - 1$ and r contain z's in at least two rows. If necessary by an interchange of the $r - 2$ first rows and the corresponding interchange of columns, we can obtain that columns $r - 1$ and r will contain a z in the row $r - 2$. Hence the determinant

$$\begin{vmatrix} b_{r-2,r-1} & b_{r-2,r} \\ b_{r-1,r-1} & b_{r-1,r} \end{vmatrix}$$

will not be identically equal to zero. Because B is of rank $r - 1$, the product of this determinant and $b_{11}\; b_{22} \cdots b_{r-3,r-3}\; b_{i,r-2} = 0$ when $i \geqq r$. Hence $b_{i,r-2} = 0$ when $i \geqq r$. Carrying on this argument we finally conclude that $b_{ik} = 0$ when $i \geqq r$ and $k = 1, 2, \cdots, r$. This results contradicts the assumption that B contains z's in r different rows. The assumption that the rank of B is $r - 1$ must therefore be false, and we conclude that the rank of B is r. This completes the proof of Lemma 8.2.

Let K_m denote the matrix which we obtain from A_m when we delete the mth column (which consists of zeros only), and let K_{mi} denote the matrix which we obtain from A_{mi} when we delete the mth column. We can now state

Theorem 8.3. When Specification (2.12) is satisfied, each of the matrices K_m *satisfies the following condition for* $q = 1, 2, \cdots, r-1$: *For any set of* q *columns of* K_m *there are at least* $q+1$ *rows each of which has a non-zero element in at least one of the* q *columns.*

Proof: When Specification (2.12) is satisfied each matrix K_{mi} is of rank $r-1$. Applying Lemma 8.2 to the matrix K_{mi}, we conclude that it satisfies the condition of this lemma with r replaced by $r-1$. If there exists a K_m for which the condition of Theorem 8.3 is not satisfied, it will be possible to remove a row of K_m such that we obtain a matrix K_{mi} which does not satisfy the condition of Lemma 8.2. But we have shown that this is not possible. Hence Theorem 8.3 is proved.

By a similar reasoning the following converse of Theorem 8.3 can be proved.

Theorem 8.4. Suppose that each of the matrices K_m *satisfies the following condition for* $q = 1, 2, \cdots, r-1$. *For any set of* q *columns of* K_m *there are at least* $q+1$ *rows each of which has a non-zero element in at least one of the* q *columns. Then Specification (2.12) is satisfied everywhere in the space generated by the non-zero elements of* A, *except for a set of measure zero.*

9. *Identifiability in the complete model.*

We shall begin by stating a theorem which is similar to Theorem 6.1.

Theorem 9.1 A necessary and sufficient condition for A *to be identifiable in Model (2.1-12) is that all matrices* A^* *which are equivalent to* A *in the model defined by Specifications (2.1) through (2.10) and (2.12) can be written in the form* $A\overline{V}$, *where* $\overline{V}$ *is a square non-singular matrix which contains exactly one non-zero element in each row and exactly one non-zero element in each column.*

Algebraically the condition for identifiability of A may be expressed in the following equivalent ways.

Theorem 9.2. When D *is identifiable in Model (2.1-9), a necessary and sufficient condition for the identifiability of* A *in Model (2.1-12) is that* A *does not contain other submatrices than* A_1, A_2, $\cdots$, A_r *satisfying Conditions (7.1) through (7.3), with* A *substituted for* B.

Theorem 9.3. When D *is identifiable in Model (2.1-9), a necsary and sufficient condition for the identifiability of* A *in Model (2.1-12) is that a matrix* B = AV *(where* V *is non-singular) contains exactly* r *submatrices* B_m *satisfying Conditions (7.1) through (7.4).*

Let us call a hyperplane in $(r-1)$-dimensional space *over-determined* by a set of points lying in the hyperplane if the hyperplane is still determined by any point set which we obtain from the first point set by removing one point, in other words if none of the sets of remaining points is situated in an $(r-3)$-flat.

We shall now formulate the condition for identifiability of A in terms of the geometrical representation.

Theorem 9.2b. When D *is identifiable in Model (2.1-9), a necessary and sufficient condition for the identifiability of* A *in Model (2.1-12) is that each hyperplane which is over-determined by the test points lying in it is a coordinate hyperplane.*

Theorem 9.3b. When D *is identifiable in Model (2.1-9), a necessary and sufficient condition for the identifiability of* A *in Model (2.1-12) is that there exist exactly* r *hyperplanes which are over-determined by the test points lying in them and that these hyperplanes are linearly independent, i.e. that they do not pass through the same point.*

Specification (2.12) may be reformulated as follows:

> Each of the coordinate hyperplanes is over-determined by the test points lying in the hyperplane. (2.12b)

The proofs of the geometric formulations 9.2b and 9.3b are now obvious.

Theorems 9.2 and 9.3 may be proved by showing the equivalence between the algebraic and geometric formulations. We shall, however, give an independent algebraic proof.

Suppose first that A does not contain any submatrix besides A_1, A_2, $\cdots$, A_r which satisfies Conditions (7.1) through (7.3). Let A and $A^* = AV$ be equivalent in Model (2.1-12). Consider one column of A^*, say the first column. Since A^* satisfies Specifications (2.10) and (2.12), there must be at least r zeros in the column. Let A^*_{r+1} be the submatrix of A^* which consists of all rows of A^* which have zeros in the first column, and let A_{r+1} be the corresponding submatrix of A. Denoting the first column of V by V_1, we have

$$A_{r+1} V_1 = 0 . \tag{9.1}$$

Because A^* satisfies Specification (2.12) and A^*_{r+1} contains all the rows of A^* which have zeros in the first column, the matrix A_{r+1} must satisfy Conditions (7.1) through (7.3). Since A does not contain any submatrix besides $A_1, A_2, \cdots, A_r$ which satisfies Conditions (7.1) through (7.3), A_{r+1} must be identical with one of the matrices $A_1, A_2, \cdots, A_r$, say A_m. Hence we have $A_m V_1 = 0$. Since all elements of the mth column of A_m are zero, all elements of V_1 must be zero except v_{m1}. The same proof can be carried through for any column of V, so that each column of V contains only one non-zero element. Since V is non-singular, it must contain one and only one element in each row. V is therefore of the type $\bar{V}$ considered in Theorem 9.1.

Inversely, suppose that there exists a submatrix A_{r+1} of A which satisfies Conditions (7.1) through (7.3) and which is different from any of the matrices $A_1, A_2, \cdots, A_r$. Then there exists a column V_1 such that $A_{r+1}V_1 = 0$. The column AV_1 will be different from any multiple of any of the columns of A, since its zeros are in different places. Hence the column V_1 must contain at least two non-zero elements. Let us now form a matrix V, taking V_1 as the first column and letting each of the other columns contain exactly one non-zero element, in such a way that V is non-singular. Then $A^* = AV$ and A are equivalent in Model (2.1-12), and the transformation matrix V is not of the type $\bar{V}$. This proves Theorem 9.2.

When $A_1, A_2, \cdots, A_r$ are the only submatrices of A satisfying Conditions (7.1) through (7.3), the matrix B will contain only r submatrices satisfying Conditions (7.1) through (7.4). Conversely, if there exists a submatrix A_{r+1} which satisfies Conditions (7.1) through (7.3), and which is different from any of the matrices $A_1, A_2, \cdots, A_r$, then there exists at least $r+1$ submatrices of B satisfying Conditions (7.1) through (7.4). From these remarks Theorem 9.3 follows.

Theorem 9.2 expresses the condition for the identifiability of A in terms of A itself. This theorem can therefore not be used to find out whether A is identifiable or not if we do not have *a priori* information about A sufficient to decide whether the condition is fulfilled or not. Theorem 9.3 on the other hand makes it possible to examine if the condition is fulfilled before A is determined and without any *a priori* information about A.

Let us finally consider the identifiability of Q when D and A are identifiable. Two equivalent structures must now have the same D and A. Let $T = \{A, Q, D\}$ and $T^* = \{A, Q^*, D\}$ be any two equivalent structures. Then Equations (5.5) and (5.6) must be valid with $A^* = A$. This gives $V = I$ and $Q = Q^*$. This result may be stated as follows:

Theorem 9.4. When D *and* A *are identifiable* Q *is also identifiable.*

10. Remarks on the determination of r

We have seen that r is always identifiable in a model containing Specifications (2.1) through (2.9). Hence r can always be determined if M is known. But the actual determination of r when the population covariance matrix is known is difficult. A still more difficult problem is the estimation of r when only a sample is known.

I shall mention some results which may be useful in determining r, although they apply only to special cases.

Theorem 10.1. Necessary and sufficient for r *to be equal to one is that both of the following conditions are satisfied*:

(a) *All tetrad differences are zero, i.e.,*

$$\begin{vmatrix} m_{ij} & m_{il} \\ m_{kj} & m_{kl} \end{vmatrix} = 0$$

when i, j, k, *and* l *are all different.*

(b)

$$0 \leqq \frac{m_{ij}\, m_{jk}}{m_{ik}} \leqq m_{jj}$$

for all sets of values of i, j, *and* k *which are all different.*

Condition (a) is due to Spearman and is well known. It is usually presented as the only condition, ignoring Condition (b). That Condition (b) is necessary follows from Spearman's formula,

$$m_{jj} - d_j = \frac{m_{ij}\, m_{jk}}{m_{ik}}, \qquad (10.1)$$

and the obvious fact that we must have

$$0 \leqq m_{jj} - d_j \leqq m_{jj} . \qquad (10.2)$$

We shall show that the conditions are also sufficient. When Condition (a) is satisfied and D is determined by (10.1), the matrix $M - D$ is of rank 1. When Condition (b) is satisfied, (10.2) holds good. This implies that the diagonal elements of $M - D$ are non-negative, and since all principal minors of order two or higher order are zero, $M - D$ is positive semidefinite. From (10.2) also follows that D satisfies Condition (3.3). This proves the sufficiency of the conditions.

If $m_{ij} = 0$, either all non-diagonal elements in the ith row and column or all non-diagonal elements in the jth row and column of M must be zero. This contradicts Specification (2.8). Hence all elements of M must be different from zero when $r = 1$.

Condition (a) may also be put in the form that if we ignore the diagonal elements, all rows of M are proportional. It is not necessary to examine whether Condition (b) is fulfilled for all sets of values of i, j, and k. For each value of j it is sufficient to consider one i and one k, because that is sufficient to ensure that (10.2) is satisfied. Hence we get a modified form of Theorem 10.1.

Theorem 10.2. Necessary and sufficient for r *to be equal to one is that each of the following conditions is fulfilled.*

(*a*) *Each element of* M *is different from zero.*

(*b*) *Excepting the diagonal elements, all rows of* M *are proportional.*

(*c*) *The inequality*

$$0 < \frac{m_{ij}\, m_{jk}}{m_{ik}} \leqq m_{jj}$$

is satisfied for some i *and* k *for each* j.

Consider the inverse M^{-1} of the covariance matrix M. Let us change the sign of each row in M^{-1} which has a negative element in the first column. Next let us change the sign of each column which has a negative element in the first row. If after these sign changes all elements of the resulting matrix are positive, we shall say that the matrix M^{-1} has *compatible signs.* Using this term we shall state the condition for $r \leqq n - 2$.

Theorem 10.3 A necessary and sufficient condition for r *to be less than or equal to* n — 2 *is that the inverse of the covariance matrix*

M *has signs which are not compatible.*

This theorem is an immediate corollary of a theorem presented in one of the writer's previous papers. (**9**, Theorem 14). From the same theorem we immediately get a necessary condition that $r \leqq s$, where $s \leqq n-2$. This condition is far from sufficient when $s < n-2$.

Theorem 10.4. A necessary condition for r *to be less than or equal to a number* s *which is less than* n — 1 *is that the inverse of each* $(s+2)$-rowed *principal minor of* M *has signs which are not compatible.*

As a special case of Theorem 10.3 we get a necessary and sufficient condition for $r \leqq 2$ when $n = 4$. This case was previously considered by Kelley (**5**, 52), Hotelling (**5**, 54) and Wilson and Worcester (**13**). A necessary condition for $r \leqq 2$ when $n = 5$ was found by Kelley (**5**, 58). A necessary and sufficient condition for the same case was found by Wilson and Worcester (**14**).

Other results have been obtained by A. A. Albert (**1**, **2**). For a given M let ρ be the minimum rank of $M - D$ when we require D to be diagonal, but do not impose the conditions that D shall be non-negative and $M - D$ positive semidefinite. Let ρ' be the highest order of a minor of M which does not contain any diagonal element of M. Albert gives necessary and sufficient conditions for the equality $\rho = \rho'$ and gives a method for determining D in the case when this equality is fulfilled. The solution arrived at need not necessarily satisfy Conditions (3.3) and (3.4).

11. Remarks on the specification defining r

Specification (2.9) is a mechanical way of defining r, and at first sight it looks impossible that an r defined in this way can represent any psychological reality. This impression is strengthened by Theorem 3.3 which gives the following immediate corollary: If there exists a structure in Model (2.1-8) with non-singular D in the case when r has its minimum value s, there exist equivalent structures in Model (2.1-8) where $r = s+1, s+2, \cdots, n-1, n$.

Nevertheless the data may be such that they give us a reason to believe that the value of r defined by Specification (2.9) does express some kind of underlying reality. In order to make this clear let us consider the following experiment:

Suppose that a person X constructs a structure in Model (2.1-12) in the following way. He chooses Q—I, and constructs a matrix

A with n rows and r_0 columns, where

$$r_0 \leqq r_n = \tfrac{1}{2}\,(2n + 1 - \sqrt{8n + 1}).$$

He chooses some of the elements in A equal to zero such that the conditions of Theorem 8.2 are fulfilled when the other elements are different from zero. He chooses each of the other elements a_{jm} by a random drawing from a parent population, for instance a population with a rectangular distribution between the limits a_{jm} and $\bar{a}_{jm}$. In practice the parent population will have to be discrete. Let us suppose all possible values of a_{ij} in the population are represented by all decimal fractions between the two limits which have a given number of decimal places. We shall suppose that the number of possible values of each a_{ij} is large, so that the probability of one particular value of a_{ij} is small.

Thereafter X determines the elements d_i in the same way by random drawings from a parent population. There will be a small probability that the structure obtained in this way will not satisfy the Specifications of Model (2.1-12). If this should happen X might construct another structure or modify the structure he already has constructed. Finally X computes the matrix M from equation (2.14) and rounds off to a certain number of decimal places.

He then presents the matrix M to another person Y. He informs Y about the method used in obtaining M, but does not give him any information about the numerical values of the elements of A and D or the value of r_0. Y now tries to find r_0, using Specifications (2.1) through (2.9), and taking into account that the equations need not be exactly fulfilled because of rounding-off errors. The question now is: Will Y find $r = r_0$, or will he find a smaller value of r?

When $r = r_0$, the d_i will satisfy $\frac{1}{2}\,(n - r_0)(n - r_0 + 1)$ equations. If $r = r_0 - 1$, the d_i will satisfy $\frac{1}{2}\,(n - r_0 + 1)(n - r_0 + 2)$ equations; i.e., in addition to the $\frac{1}{2}\,(n - r_0)(n - r_0 + 1)$ equations which must be satisfied because of the way in which the matrix M has been obtained, there will exist $n - r_0 + 1$ additional equations between the parameters which must happen by chance to be satisfied, and there is only a small probability that this will happen. In other words, there is only a small probability that Y will not find $r = r_0$.

Suppose next that X does not give Y the population covariance matrix, but a sample from a normal probability distribution with covariance matrix M. If the sample is large enough there will still be only a small probability that Y will not estimate correctly the value r_0.

From the standpoint of the person Y the argument will be slightly different. When he finds a value r by assuming Specifications (2.1) through (2.9), he will ask whether this is the true value r_0 or if r_0 is larger than r. If $r \geqq r_n$, Y will not have any confidence in the result, but if $r < r_n$, the sample which he has got would represent an improbable event if $r_0 > r$ and the sample is large enough.

Thus we can only have faith in the results of a factor analysis if the parameters are *overidentified,* i.e., if the number of independent equations, which determine the parameters in terms of the probability distribution of the observed variables, is greater than the number of parameters. (Cf. **12**, 294).

In Section 7 we found that the central part of Model (2.1-9) cannot be tested. However, if we add the assumption that $r < r_n$, the model can be tested; and if we for sufficiently large samples regularly find that $r < r_n$, this will represent a verification of the model.

It should be pointed out that we have throughout this paper assumed a fixed battery of tests and a fixed population of individuals tested. If we consider the possibility of a partial change of the test battery and a change of the population, both the question of the possibility of testing the model and the question of identifiability of the parameters come in a new light. (Cf. **12**, Chapter 16).

REFERENCES

1. Albert, A. A. The matrices of factor analysis. *Proc. nat. Acad. Sci., Wash.*, 1944, 30, 90-95.
2. Albert, A. A. The minimum rank of a correlation matrix. *Proc. nat. Acad. Sci., Wash.*, 1944, 30, 144-146.
3. Haavelmo, T. The probability approach in econometrics. Supplement to *Econometrica*, 1944, 12.
4. Hurwicz, L. Generalization of the concept of identification. *Cowles Commission Monograph No. 10. Statistical inference in dynamic economic models.* New York: John Wiley and Sons, 1949.
5. Kelley, T. L. Crossroads in the mind of man. Stanford University, Calif.: Stanford Univ. Press, 1928.
6. Koopmans, T. C., Rubin, H., and Leipnik, R. B. Measuring the equation systems of dynamic economics. *Cowles Commission Monograph No. 10.* New York: John Wiley and Sons, 1949.
7. Koopmans, T. C., and Reiersøl, O. The identification of structural characteristics. To be published.
8. Ledermann, W. On the rank of the reduced correlational matrix in multiple factor analysis. *Psychometrika,* 1937, 2, 85-93.

9. Reiersøl, O. Confluence analysis by means of lag moments and other methods of confluence analysis, *Econometrica*, 1941, **9**, 1-24.
10. Thomson, G. H. The proof or disproof of the existence of general ability. *Brit. J. Psychol.*, 1919, **9**, 323-336.
11. Thurstone, L. L. The vectors of mind. Chicago: Univ. Chicago Press, 1935.
12. Thurstone, L. L. Multiple-factor analysis. Chicago: Univ. Chicago Press, 1947.
13. Wilson, E. B., and Worcester, J. The resolution of four tests. *Proc. nat. Acad. Sci., Wash.*, 1934, **20**, 189-192.
14. Wilson, E. B., and Worcester, J. The resolution of tests into two general factors. *Proc. nat. Acad. Sci., Wash.*, 1938, **25**, 20-25.
15. Wilson, E. B., and Worcester, J. The resolution of six tests into three general factors. *Proc. nat. Acad. Sci., Wash.*, 1939, **25**, 73-77.